Авторы:

Колбасина Ирина Валерьевна

аспирант

Попов Никита Алексеевич

аспирант

Научный руководитель:

Егорушкин Олег Игоревич

соискатель, старший преподаватель

ФГБОУ ВО «Сибирский государственный университет

науки и технологий им. академика М.Ф. Решетнева»

г. Красноярск, Красноярский край

О МЕТОДЕ СИНТАКСИЧЕСКОГО АНАЛИЗА КОНТЕКСТНО-СВОБОДНЫХ ЯЗЫКОВ

Аннотация: *в статье разрабатывается подход к синтаксическому анализу контекстно-свободных языков, которые являются решением систем некоммутативных полиномиальных уравнений. Результаты имеют приложение в теории формальных языков.*

Ключевые слова: *контекстно-свободные языки, полиномиальные уравнения, формальный степенной ряд, синтаксический анализ.*

В связи с приложениями в теории языков программирования целесообразно рассмотреть систему полиномиальных уравнений

$$P_j(z,x)=0,\ j=1,...,n, \tag{1}$$

где $z=(z_1,...,z_n), x=(x_1,...,x_m)$ – переменные из кольца с некоммутативной операцией умножения и коммутативной операцией сложения; для них определена также коммутативная операция умножения на комплексные числа; при этом система (1) решается относительно переменных $z_1,...,z_n$ в виде формальных степенных рядов (ФСР) от переменных $x_1,...,x_m$. А именно, в приложениях система (1) рассматривается как грамматика над терминальными символами

$x_1,...,x_m$ – словами языка, и нетерминальными символами $z_1,...,z_n$, необходимыми для задания грамматики [1–4].

Одной из важных проблем, связанных с разработкой систем и языков программирования, является проблема синтаксического анализа программ. Как было отмечено выше, большинство языков программирования является кс-языками, которые можно представить в виде ФСР, поэтому каждая программа, написанная на языке программирования, может рассматриваться как моном соответствующего ФСР. В связи с этим рассмотрим проблему синтаксического анализа мономов кс-языка.

Для того, чтобы сформулировать эту проблему, рассмотрим грамматику кс-языка, которая является множеством правил подстановки

$$z_j \to q_{j1}(z,x),...,z_j \to q_{jp_j}(z,x),\ j=1,...,n, \tag{2}$$

где $q_{jk}(z,x)$ является мономом от некоммутативных символьных переменных с числовым коэффициентом, равным единице. Правила подстановки можно применять к начальному символу z_1, а затем к другим мономам в любом порядке неограниченное число раз, что позволяют выводить новые правильные мономы, образующие кс-язык.

Итак, проблема синтаксического анализа мономов состоит в том, чтобы определить, принадлежит ли моном данному кс-языку, т.е. может ли быть получен из начального символа z_1 при помощи правил подстановки (2), а также установить, какие правила подстановки и сколько раз использовались при выводе этого монома; при этом порядок использования правил подстановки не имеет значения.

Метод беступикового синтаксического анализа, основанный на использовании мономиальных меток, был предложен в статье [1]. В соответствии с этим методом, каждое правило подстановки $z_j \to q_{jk}(z,x)$ заменяется правилом $z_j \to t_{jk}q_{jk}(z,x)$, имеющим мономиальную метку t_{jk}, которая является символом из расширенного алфавита, и для новых правил вывода рассматривается соответствующая система уравнений Хомского-Шутценберже:

$$z_j = Q_j^*(z,x,t) =: t_{j1}q_{j1}(z,x) + \ldots + t_{jp_j}q_{jp_j}(z,x),\ j = 1,\ldots,n. \qquad (3)$$

Рассмотрим метод мономиальных меток, который позволяет провести синтаксический анализ монома v от терминальных символов $x_1,\ldots,x_m$. Итерации метода последовательных приближений для системы уравнений (3) дают многочлены возрастающей степени относительно символов $x_1,\ldots,x_m,\ t_{11},t_{12},\ldots,t_{np_n}$; при этом мономы степени не выше $\deg_x(v)$ относительно символов $x_1,\ldots,x_m$ после конечного числа итераций стабилизируются, не меняясь при последующих итерациях. Таким образом, можно получить начальные члены решения системы (7) в виде ФСР до любой, сколь угодно высокой степени, в том числе члены ФСР, представляющего первую компоненту этого решения:

$$z_1 = z_1^*(x,t) = \sum_{i=0}^{\infty} < z_1^*, w_i > w_i. \qquad (4)$$

Синтаксический анализ монома v кс-языка $z_1(x)$ можно провести следующим образом. Считывая мономы степени $\deg_x(v)$ относительно символов $x_1,\ldots,x_m$ и пропуская символы t_{jk}, можно установить, есть ли среди них моном v, а значит, можно ли вывести его с помощью системы продукций (2). При этом каждая мономиальная метка t_{jk}, содержащаяся в таком мономе, показывает, что при его выводе использовалось правило $z_j \to t_{jk}q_{jk}(z,x)$. В самом деле, из системы уравнений (3) и метода последовательных приближений нетрудно видеть, что, применяя это правило вывода к моному, мы умножаем его слева на символ t_{jk}. Следовательно, мономиальные метки монома решают проблему его синтаксического анализа, показывая, какие правила вывода кс-языка и сколько раз использовались при выводе этого монома, с точностью до порядка их применения.

Таким образом, имеет место следующая теорема [1].

Теорема. Метод мономиальных меток позволяет провести за конечное число шагов беступиковый синтаксический анализ любого монома (программы) кс-языка, заданного грамматикой (2).

Недостатком метода мономиальных меток является большое число громоздких итераций метода последовательных приближений, необходимых для получения ФСР, который представляет кс-язык, причём это число возрастает вместе с ростом степени монома, который анализируется. В связи с этим актуально найти другой путь для получения мономиальных меток некоммутативного монома.

Кроме того, используя систему скобок при записи правил вывода, можно установить порядок их применения в случаях, когда это имеет смысл.

Список литературы

1. Сафонов К.В. О синтаксическом анализе и проблеме В.М. Глушкова распознавания контекстно-свободных языков Хомского / К.В. Сафонов, О.И. Егорушкин // Вестник Томского государственного университета. – 2006. – №17. – С. 63–67.

2. Safonov K.V. On conditions for the sum of a power series to be algebraic and rational // Mathematical Notes. – 1987. – №3 (41). – P. 185–189.

3. Safonov K.V. On Power Series of Algebraic and Rational functions in Cn // Journal of Math. Analysis and Applications. – 2000. – V. 243. – P. 261–277.

4. Salomaa A. Automata-Theoretics Aspects of Formal Power Series / A. Salomaa, M. Soitolla. – N.-Y.: Springer Verlag, 1978. – 171 p.